MICROCOPY RESOLUTION TEST CHART
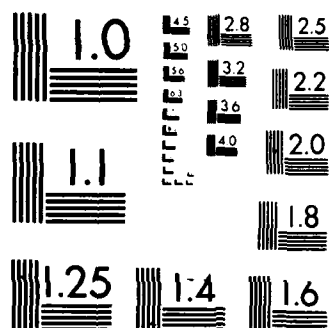
NATIONAL BUREAU OF STANDARDS-1963-A

**LABORATORY FOR COMPUTER SCIENCE**

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

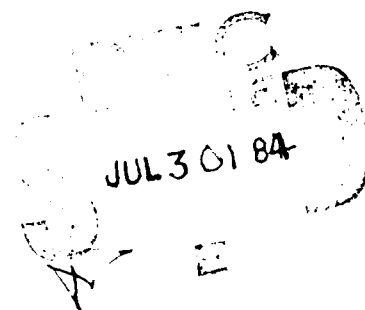# THREE-DIMENSIONAL CIRCUIT LAYOUTS

Thomas Leighton

Arnold Rosenberg

June 1984

JUL 3 01 84

545 TECHNOLOGY SQUARE, CAMBRIDGE, MASSACHUSETTS 02139

84 07 23 150

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| **1. REPORT NUMBER** <br> MIT/LCS/TM-262 | **2. GOVT ACCESSION NO.** <br> A143 430 | **3. RECIPIENT'S CATALOG NUMBER** |
| **4. TITLE** *(and Subtitle)* <br> Three-Dimensional Circuit Layouts | | **5. TYPE OF REPORT & PERIOD COVERED** <br> Interim Research |
| | | **6. PERFORMING ORG. REPORT NUMBER** <br> MIT/LCS/TM-262 |
| **7. AUTHOR(s)** <br> Tom Leighton & <br> Arnold L. Rosenberg | | **8. CONTRACT OR GRANT NUMBER(s)** <br> DARPA/DOD C <br> N00014-80-0622 |
| **9. PERFORMING ORGANIZATION NAME AND ADDRESS** <br> MIT Laboratory for Computer Science <br> 545 Technology Square <br> Cambridge, MA 02139 | | **10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS** |
| **11. CONTROLLING OFFICE NAME AND ADDRESS** <br> DARPA/DOD <br> 1400 Wilson Boulevard <br> Arlington, VA 22209 | | **12. REPORT DATE** <br> June 1984 |
| | | **13. NUMBER OF PAGES** <br> 33 |
| **14. MONITORING AGENCY NAME & ADDRESS**(*if different from Controlling Office*) <br> ONR/Department of the Navy <br> Information Systems Program <br> Arlington, VA 22217 | | **15. SECURITY CLASS.** *(of this report)* <br> Unclassified |
| | | **15a. DECLASSIFICATION/DOWNGRADING SCHEDULE** |

**16. DISTRIBUTION STATEMENT** *(of this Report)*

Approved for public release; distribution unlimited.

**17. DISTRIBUTION STATEMENT** *(of the abstract entered in Block 20, if different from Report)*

Unlimited.

**18. SUPPLEMENTARY NOTES**

**19. KEY WORDS** *(Continue on reverse side if necessary and identify by block number)*

area, bifurcator, decomposition tree, multiple layers, placement and routing, three-dimensional layouts, very large scale integration, volume, wire length.

**20. ABSTRACT** *(Continue on reverse side if necessary and identify by block number)*

Recent advances in fabrication technology have rendered imminent the fabrication of multilayer chips, wafers and circuit boards. In this paper, we examine the savings in material and communication time afforded by the development of three-dimensional technology. In particular, we derive close upper and lower bounds on the volume and maximum wire length with which circuits can be realized in a multilayer medium. For example, we find that the

DD FORM 1473  1 JAN 73  EDITION OF 1 NOV 68 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

smallest volume of any three-dimensional layout of an N-device circuit is no more than (roughly) $(AN)^{\frac{1}{2}}$, where A is the smallest area of any two-dimensional layout of the circuit. We also show how to efficiently transform a two-dimensional layout of area A and maximum wire length L into a three-dimensional layout of volume (roughly)V=A/H and maximum wire length L*=L/H for moderate numbers of layers H.   Two noteworthy features of the study are:

1)   that, within logarithmic factors, the indicated savings can be realized with layouts that use the third dimension only for interconnect; and
2)   that the indicated savings can be realized algorithmically: we present polynomial-time algorithms that transform a given two-dimensional layout into a more efficient three-dimensional layout.

# Three-Dimensional Circuit Layouts

Tom Leighton

Mathematics Department and
Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Arnold L. Rosenberg

Department of Computer Science
Duke University
Durham, North Carolina 27706
and
Microelectronics Center of North Carolina
Research Triangle Park, North Carolina 27709

**Abstract:** Recent advances in fabrication technology have rendered imminent the fabrication of multilayer chips, wafers and circuit boards. In this paper, we examine the savings in material and communication time afforded by the development of three-dimensional technology. In particular, we derive close upper and lower bounds on the volume and maximum wire length with which circuits can be realized in a multilayer medium. For example, we find that the smallest volume of any three-dimensional layout of an $N$-device circuit is no more than (roughly) $(AN)^{1/2}$, where $A$ is the smallest area of any two-dimensional layout of the circuit. We also show how to efficiently transform a two-dimensional layout of area $A$ and maximum wire length $L$ into a three-dimensional layout of volume (roughly) $V = A/H$ and maximum wire length $L^* = L/H$ for moderate numbers of layers $H$. Two noteworthy features of the study are:

1) that, within logarithmic factors, the indicated savings can be realized with layouts that use the third dimension only for interconnect; and

2) that the indicated savings can be realized algorithmically: we present polynomial-time algorithms that transform a given two-dimensional layout into a more efficient three-dimensional layout.

**Key Words:** area, bifurcator, decomposition tree, multiple layers, placement and routing, three-dimensional layouts, very large scale integration, volume, wire length.

# THREE-DIMENSIONAL CIRCUIT LAYOUTS

## 1. INTRODUCTION

Recent advances in fabrication technology [4-8,10-12,17-19,22,28] have allowed circuit and system designers to begin using the third dimension in realizing their designs. Multilayer packages with impressive performance have been fabricated [9,10,20], and there has been extensive research toward the goal of three-dimensional chips [8,12,17-19,28]. The rapid rate of progress in VLSI technology suggests that multilayer chips and packages will be commonplace in the not-distant future. Indeed, the president of Texas Instruments (quoted in [8]) predicts the production of three-dimensional chips by the end of the decade.

One expects (at least) *three benefits* to accrue from the use of the third dimension in circuit realization. First, wire-routing should become easier and more systematic. Next, since one can avoid obstacles by using the third dimension, runs of wire should be shorter, at least in the worst case. Finally, since avoiding obstacles in a two-dimensional environment can require area-consuming circuitous routing of wires, one would expect savings in material: the Volume of a three-dimensional realization of a circuit should be less than the Area of any two-dimensional realization of the circuit. In order to realize these expected benefits, we must develop effective techniques for devising and analyzing multilayer circuit layouts. Such is the goal of this paper: we develop and analyze an algorithmic strategy for laying out VLSI circuits -- viewed here as undirected graphs -- in three-dimensional chips -- viewed here as three-dimensional grids.

Our notion of the layout of a circuit follows the two-dimensional framework of [2,13,14,16,26,27], as adapted for the third dimension in [23,24]: circuits are

-1-

undirected graphs whose vertices correspond to active devices (transistors, gates, etc.) and whose edges correspond to wires connecting these devices. The media in which the circuits are to be realized are (two- or three-dimensional) rectangular grids. A circuit layout is an edge-disjoint embedding of the circuit-graph in the grid. Two models have been proposed for studying three-dimensional VLSI [23,24]. The first, *one-active-layer*, model requires that all active devices be placed on a designated layer of the chip. The second, unrestricted, *many-active-layer*, model allows devices to be placed arbitrarily throughout the chip. It is clear how these two possibilities manifest themselves in our formal setting. Although the many-active-layer model affords one more flexibility when laying out one's circuits, it places significantly more stringent demands on the fabrication technology; cf. [24]. There is thus a tradeoff between the cost of fabricating a chip with multiple layers of devices and the savings (in terms of Volume and maximum wire run) resulting from the increased layout flexibility. One of our more surprising results here is that, at least within our abstract framework, many-active-layer layouts are little or no more efficient than one-active-layer layouts when the number of layers is relatively small: either mode of using the third dimension affords one appreciable savings over any two-dimensional layout. Additionally, we show that multiple layers are effective in reducing Volume and maximum wire run only up to a certain point, after which they are wasteful. Although these results are definitive only for the theoretical model our analysis is based on, they suggest strongly that VLSI chips that have a higher (and costlier) degree of sophistication (in terms of number of layers and placement of devices) may not be more efficient for many applications than significantly more modest chips.

Although there has been a substantial amount of work on the two-dimensional version of the layout problem, related work on the three-dimensional problem has largely been confined to one of:

* the study of routing in the presence of a few extra layers [3,9,21];

-2-

* the study of optimal multilayer layouts for a few special networks [20,23,29];

* the study of optimal multilayer layouts for the class of "hardest-to-realize" networks [23,24].

Notable among the results in these papers, for our purposes, is the use in [23,24] of optimal three-dimensional layouts of the N-input Benes permutation network [1] to prove:

Every small-degree N-vertex graph can be laid out in a three-dimensional grid with Volume $O(N^{3/2})$ and wire-length $O(N^{1/2})$.

(There exist graphs that do not admit any more compact layout; for such graphs, these bounds contrast with the lower bounds of Area $\Omega(N^2)$ and wire-length $\Omega(N)$ [12,26] in the two-dimensional case. In effect, the contribution of the present paper is to generalize the specialized three-dimensional results of Rosenberg and Preparata (among others) to a level of generality comparable to the two-dimensional work of Bhatt, Leighton, Leiserson, Thompson, and Valiant (among others). Perhaps the most important contribution of this paper is an algorithm that transforms a two-dimensional circuit layout of Area A and maximum wire run L into a three-dimensional layout of the circuit that is within logarithmic factors of Volume A/H and maximum wire run L/H, for moderate values of H. The layouts produced are close to optimal in the sense that using H layers rather than just one layer (which is how the two-dimensional case is viewed in our formal framework) can never improve Area or wire-length by a smaller factor than 1/H. Certain special situations wherein the logarithmic factors can be avoided are described in [15], wherein is also a special case of our algorithm.

The remainder of the paper is divided into four sections. In Section 2 we review basic definitions and cite work on two-dimensional layouts that is relevant to our study. Sections 3 and 4 are devoted to the development and analysis of our three-dimensional layout strategy, with particular attention paid to issues of Volume and

-3-

maximum wire run. We conclude in Section 5 with some remarks on the implications of our work.

## 2. PRELIMINARIES

*Underlying Assumptions.* The formal framework of our study carries with it certain implicit assumptions:

1.  Our associating circuits with graphs limits our study to circuits with two-point nets.

2.  Our associating chips with grids limits our circuits to having small vertex-degrees.

3.  Our adherence to the models of VLSI layout theory renders the vertices of our circuits as unit-side squares or cubes.

4.  Our method of extending the two-dimensional model assumes isometry in all dimensions: a unit of height is eqivalent to a unit of width.

It is worthwhile placing these assumptions in perspective.

1.  The restriction to two-point nets is a significant one: although extending our results to circuits with three- or four-point nets is not difficult, extending the results to circuits with arbitrary multipoint nets remains an inviting challenge.

2.  Techniques that are now standard can be used to generalize our results to include circuits with high vertex-degrees, but the associated analysis is technically somewhat more complicated.

3.  Restricting attention to unit-side devices is a purely clerical device; extending the analysis to any uniform-size devices should present no problem [2,13].

4.  Aside from clerical simplification, the isometry assumption acknowledges the potential problem of cross-talk between parallel runs of wire [24].

*The Formal Framework.* An undirected graph comprises a finite set V of *vertices* and a set of two-element subsets of V, called *edges*. We say that the edge {u,v} is

*incident* to vertices u and v. The *degree* of the vertex v is the number of edges incident to v; the *degree* $D(G)$ of G is the largest degree of any of its vertices.

The W×L *planar grid* is the graph whose vertex-set is the set of pairs $[W]\times[L]$ and whose edges connect vertices $<a,b>$ and $<c,d>$ just when $|a-c| + |b-d| = 1$. (Here and throughout, $[n]$ denotes the set $[n] = \{1, 2, ..., n\}$.) The H×W×L *solid grid* is the graph whose vertex-set is the set of triples $[H]\times[W]\times[L]$ and whose edges connect vertices $<a,b,c>$ and $<d,e,f>$ just when $|a-d| + |b-e| + |c-f| = 1$.

An *embedding* or *layout* of the graph G in the grid $\Gamma$ (solid or planar) is a one-to-one association of the vertices of G with vertices of $\Gamma$, together with a one-to-one association $\alpha$ of the edges of G with edge-disjoint paths in $\Gamma$. An embedding in a solid grid $\Gamma$ of dimensions H×W×L is a *one-active-layer* embedding if it associates all vertices of G with vertices of $\Gamma$ of the form $<i_0, j, k>$ for some fixed layer $i_0$ in $[H]$.

We gauge the cost of an embedding of a graph in a grid in terms of the amount of material consumed by the embedding (Area in the two-dimensional case and Volume in the three-dimensional case), and in terms of the maximum length of any run of wire that does not encounter a device.

The *Volume* (resp., *Area*) *of an embedding* of the graph G in a solid (resp., planar) grid $\Gamma$ is the product of the dimensions of $\Gamma$. The *Volume* (resp., *Area*) *of the graph* G, $VOL(G)$ (resp., $AREA(G)$), is the minimum Volume (resp., Area) of any embedding of G in a solid (resp., planar) grid. The *one-active-layer Volume of the graph* G, $VOL_{1-AL}(G)$, is the minimum Volume of any one-active-layer embedding of G in a solid grid. When we relativize either $VOL(G)$ or $VOL_{1-AL}(G)$ with the integer parameter H, as in $VOL(G;H)$ or $VOL_{1-AL}(G;H)$, it is to be understood that the volume minimization is done over all H-layer embeddings (of the appropriate kind).

Say that we are considering an embedding of the graph G in a grid, with the (graph edge)-(grid-path) association $\alpha$. The *wire-length of the embedding* is the maximum length of any path $\alpha(e)$ over all edges e of G. This corresponds informally to the

-5-

length of the longest run of wire that does not encounter a device. The *solid* (resp., *planar*) *wire –length of the graph* G, $WL_3(G)$ (resp., $WL_2(G)$), is the minimum wire-length of any embedding of G in a solid (resp., planar) grid. The *one-active-layer wire-length of the graph* G, $WL_{1-AL}(G)$, is the minimum wire-length of any one-active-layer embedding of G in a solid grid. As before, relativization of these measures with the integer parameter H, as in $WL_3(G;H)$ or $WL_{1-AL}(G;H)$, restricts the indicated minimization to H-layer embeddings of the appropriate kind.

Leiserson [16] and Valiant [27] showed that the "decomposition structure" of a graph could be exploited in order to find an efficient two-dimensional layout of the graph. Leighton [14] and Thompson [26] proved that the Leiserson-Valiant strategy could not be improved in general, though it often produced layouts that could be dramatically improved. Bhatt and Leighton [2,13] significantly improved the layout strategy by recasting its framework. Specifically, they reformulated the underlying notion of the "decomposition structure" of a graph to one in which the Leiserson-Valiant strategy yielded layouts that were provably good, in the sense of being within logarithmic factors of optimal, for any graph. One of the central ideas in the Bhatt-Leighton framework is that of a *decomposition tree* for a graph. The graph G has an $(F_0, F_1, ..., F_r)$-*decomposition tree* if G can be decomposed into two subgraphs $G_0$ and $G_1$ by removing at most $F_0$ edges from G; each of $G_0$ and $G_1$ can be decomposed into two subgraphs by removing at most $F_1$ edges from each; and so on, until each subgraph produced by the decomposition is either empty or an isolated vertex. See Fig. 1.

Decomposition trees for which the $F_i$ decrease at a uniform rate are of particular importance to us. A graph that has an $(F, F/\rho, F/\rho^2, ..., 1)$-decomposition tree for some real $\rho > 1$ is said to have an $(F, \rho)$-*bifurcator* or, equivalently, a $\rho$-*bifurcator of size* F. Since the decomposition tree of an N-vertex graph must have at least log N levels, it is clear that $F \geq N^{\log \rho}$. (Unless otherwise indicated, all logarithms are to the base 2.) For convenience, we shall also assume that $F < N/2$ for all graphs.
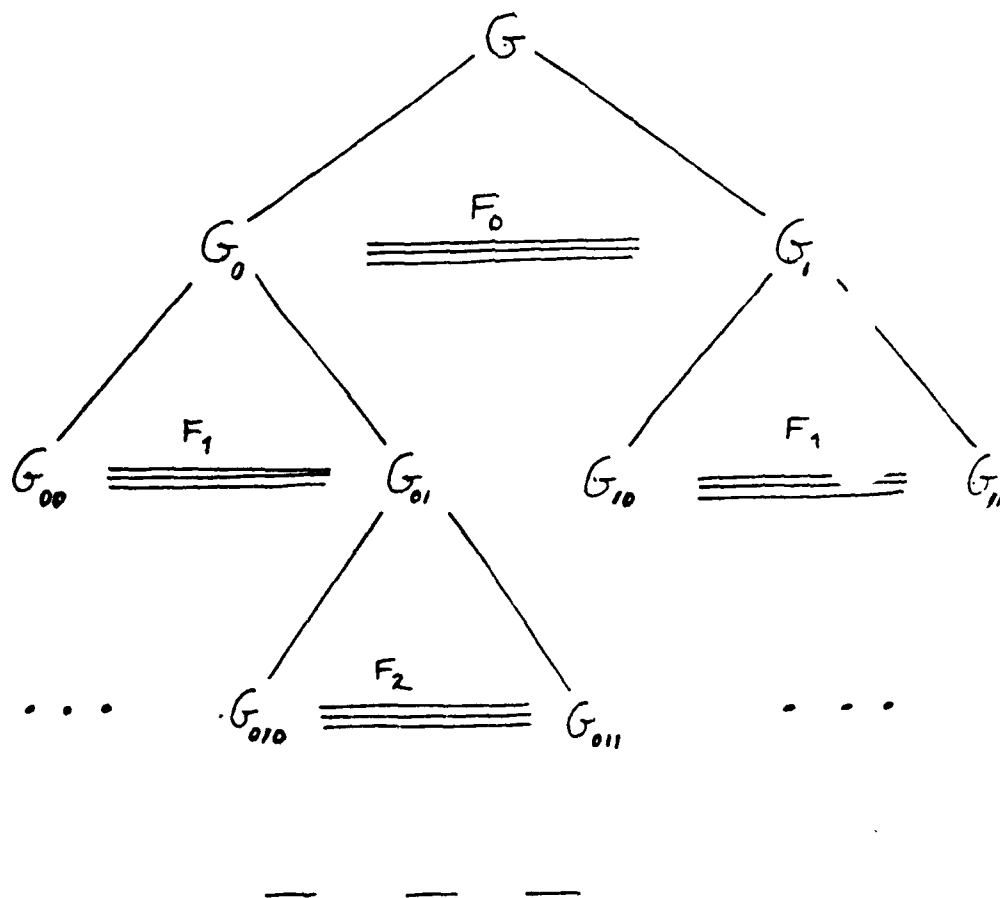
-6-

**Figure 1.** An $(F_0, F_1, \cdots, F_r)$-decomposition tree.

Returning to the issue of efficient two-dimensional layouts, Bhatt and Leighton proved that finding a small $2^{1/2}$-bifurcator for the graph to be laid out was the entire story, in the sense of the following result.

*Theorem 2.1.* [2,13] Let F be the size of the smallest $2^{1/2}$-bifurcator of the N-vertex graph G. Then

$$F^2 \leq AREA(G) \leq (const)F^2\log^2(N/F).$$

and

$$(const)F^2/N \leq WL_2(G) \leq (const)F \log(N/F)/\log\log(N/F).$$

A key step in the proof of these bounds is the demonstration that an arbitrary decomposition tree can be *fully balanced* at little or no cost, in the sense that

(1) each graph $G_i$ in the tree is split into two equal-size subgraphs, $G_{i0}$ and $G_{i1}$;

and

(2) the number of edges of G having precisely one end in the (arbitrary) tree-vertex/subgraph $G_{ia}$ of G is at most a small fixed multiple of the number of edges leaving $G_{ia}$ to go to its brother subgraph $G_{i\bar{a}}$.

The notion "fully balanced" applies to $\rho$-bifurcators in the obvious way. Bhatt and Leighton prove the following basic result, via a polynomial-time algorithm for constructing a fully balanced bifurcator from a given arbitrary one.

*Lemma 2.2.* [2,13] There is a fixed constant c>0 such that, if the graph G has a $\rho$-bifurcator of size F, then it has a fully balanced $\rho$-bifurcator of size cF.

Lemma 2.2 guarantees that any graph with an $(F, \rho)$-bifurcator has a decomposition tree in which any subgraph $G_w$ on level i of the tree is incident to at most $cF/\rho^i$ edges of G that are not wholly contained within $G_w$. Lacking the Lemma, we would know only that at most $F/\rho^i$ edges of G linked $G_w$ to its *brother* in the decomposition tree (as opposed to *any* other subgraph at level i of the tree).

A second technical lemma is crucial to our layout strategy. A *multigraph* comprises a set V of *vertices* and a multiset M of doubleton subsets of V, called *edges*

-8-

Thus a multigraph can be viewed as a graph in which each pair of vertices can be connected by several edges. The notions of "incidence", "degree of a vertex", and "degree of a multigraph" derive immediately from the corresponding notions for graphs. An *edge-coloring* of a multigraph is a labelling of the edges of the multigraph with "colors" in such a way that edges incident to the same vertex get labelled with distinct colors. Shannon [25] showed, via an efficient algorithm for edge-coloring multigraphs, that one needs never use a lot of colors to edge-color a small-degree multigraph.

*Lemma 2.3.* [25] Any multigraph G can be edge-colored using at most $\lfloor 3D(G)/2 \rfloor$ colors. Moreover, this bound is existentially tight.

## 3. EFFICIENT THREE-DIMENSIONAL LAYOUTS

### 3.1. One-Active-Layer Layouts

We consider first the problem of embedding a graph in a three-dimensional grid in accordance with the one-active layer model, i.e., so that all of the graph's vertices reside on a single layer of the layout. We assume that we have in hand a minimal-size $(F, 2^{1/2})$-bifurcator for the graph to be laid out, as well as an associated recursive decomposition of G.

*Theorem 3.1: THE ONE-ACTIVE-LAYER LAYOUT THEOREM*

Let G be an N-vertex graph, and let F be the size of its minimum $2^{1/2}$-bifurcator.

*Height-H Layout.*

There is a constant h>0 such that, for any height H in the range

$$1 \leq H \leq h \frac{F}{N^{1/2}} \log(N/F),$$

the height-H one-active-layer layouts of G satisfy

$$\max\left( FN^{1/2}, \frac{F^2}{H} \right) \leq VOL_{1-AL}(G;H) \leq (const)\frac{F^2}{H}\log^2(N/F);$$

and

$$(const)\max\left( \frac{F}{N^{1/2}}, \frac{F^2}{HN} \right) \leq WL_{1-AL}(G;H) \leq (const)\frac{F}{H}\log(N/F).$$

*Unrestricted-Height Layout.*

The minimum-resource one _active-layer layout of G satisfies

$$FN^{1/2} \le VOL_{1-AL}(G) \le (const)FN^{1/2}log(N/F);$$

and

$$(const)\frac{F}{N^{1/2}} \le WL_{1-AL}(G) \le (const)N^{1/2};$$

moreover, the number of layers (H) that minimizes $VOL_{1-AL}$ is at most $(const)\frac{F}{N^{1/2}}log(N/F)$.

Given F and an associated recursive decomposition of G, the embeddings yielding the upper bounds can be found in time polynomial in N.

*Proof.* Let G and F be as in the statement of the Theorem.

*The Lower Bounds.*

We present two proofs that expose different aspects of the situation.

*Proof 1.* Consider an arbitrary one-active-layer layout of G, having Volume V, height H, and base area B. Let us recursively bisect this "box" across the smaller of its base dimensions, in such a way that the base area is halved with each bisection. The boxes we bisect at stage i of this recursion (we start at stage 0) have height H and base area $B/2^i$. When we bisect each of these boxes, we are severing no more than $(B/2^i)^{1/2}H$ edges of G, since the area of the cutting plane is no greater than this, and wires have unit cross-sections. This means that G has a $(B^{1/2}H, 2^{1/2})$-bifurcator. Since F is the size of G's smallest $2^{1/2}$-bifurcator, it is immediate that $F \le B^{1/2}H$ so $F^2 \le BH^2 = VH$, which yields immediately that

$$V \ge \frac{F^2}{H}.$$

Since, moreover, $B \ge N$ (since all the vertices of G lie on one layer), we conclude that
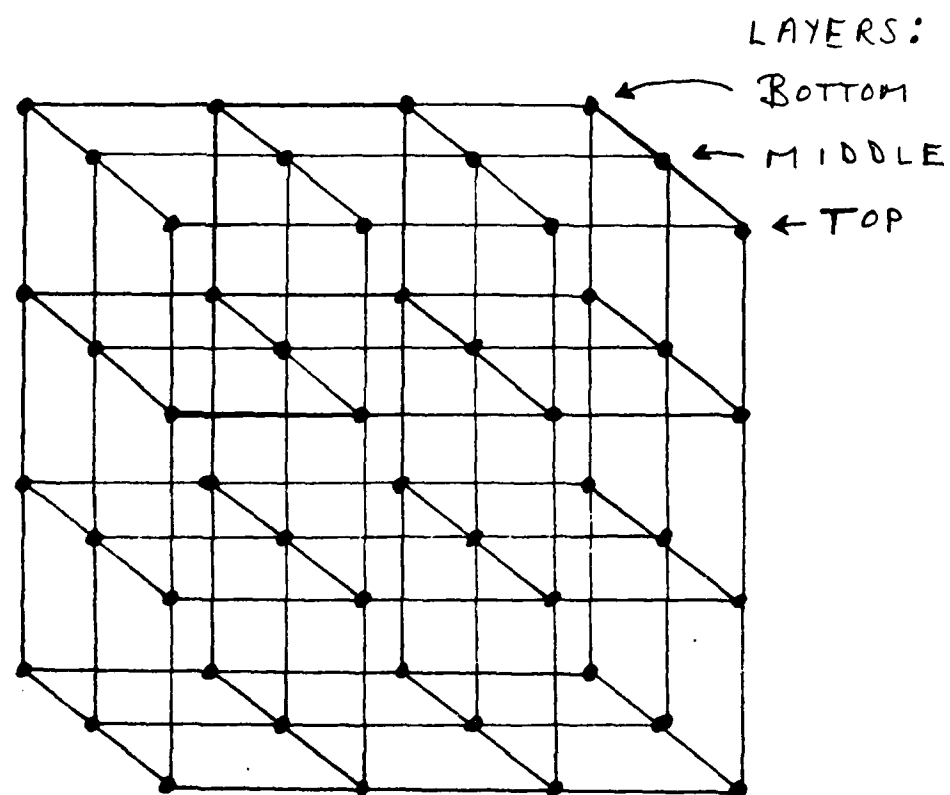
$$F \le \frac{V}{B^{1/2}} \le \frac{V}{N^{1/2}},$$

whence

**Figure 2.** The two-dimensional projection of the 3-layer 4x4 grid.

$$V \geq FN^{1/2}.$$

Since we have been looking at an arbitrary height-H one-active-layer layout of G, the lower bounds on Volume follow.

*Proof 2.* Our second, indirect, proof yields a lower bound on wire-length also.

The key step here is to transform an H-layer layout of the graph G with Volume V = BH (B being the area of the base of the layout) into a two-dimensional layout with Area BH². We shall then be able to conclude that

$$AREA(G) \leq BH^2 = VH,$$

so that

$$V \geq \frac{A}{H}.$$

Since (as before) B ≥ N, we shall also be able to conclude that

$$AREA(G) \leq BH^2 = \frac{V^2}{B} \leq \frac{V^2}{N},$$

so that

$$V \geq (AN)^{1/2}.$$

By Theorem 2.1 we know that $AREA(G) \geq F^2$, and thus

$$V \geq \max\left[FN^{1/2}, \frac{F^2}{H}\right].$$

The desired transformation is obtained by projecting the H-layer grid onto the plane, as illustrated in Fig. 2. Ignoring for the moment that the projection produces diagonal edges, it converts an HxWxL solid grid into an HWxHL planar grid. We remove the diagonal edges by rerouting the wire segments they contain. These segments are precisely the ones that run between adjacent layers of the original multilayer layout. By the rules of our model, at most one wire can pass through an intersection point that contains a wire that changes layers. Hence a wire that runs in a diagonal edge can simply be rerouted in the neighboring "right angle". (Any wire that already resides in a segment of that right angle must be electrically equivalent to the wire being rerouted.) The area of the resulting two-dimensional layout is

-12-

$$WLH^2 = BH^2,$$

as was claimed.

The same transformation yields the lower bound on wire-length: the projection maps unit-length vertical and horizontal segments into length-H segments. Unit-length segments that run between layers are transformed into segments of length 2 ($\leq$ H) when rerouted. Thus a wire of length $L_3$ in the H-layer layout is transformed into a wire of length $L_2 \leq HL_3$ in the two-dimensional layout. We can now apply Theorem 2.1 to conclude that

$$L_3 \geq (const)\frac{F^2}{NH}.$$

Since all of the vertices lie on a single layer, we know also that $L_3 \geq H/2$, which combines with the previous inequality to show that

$$L_3 \geq (const)\frac{F}{N^{1/2}}.$$

as was claimed.

Although we did not do so here, we could also show that the *average* H-layer wire-length for any N-vertex graph is at least

$$(const)max\left[\frac{F}{N^{1/2}}, \frac{F^2}{HN}\right]$$

*The Upper Bounds.*

The upper bounds are substantially more intricate to establish. Our task is lightened, though, by the fact that we can establish both the restricted-height and unrestricted-height upper bounds via a single construction, the latter bound following from the former by assigning to H its maximum allowed value. We shall, therefore, prove only the restricted-height upper bound, assuming that a legitimate target height H has been specified; this H is fixed henceforth. We establish the bound by means of a construction that recursively produces an embedding with the desired Volume for a subgraph $G_i$ of G on level i of G's decomposition tree, given the appropri-

-13-

ate embeddings of the four subgraphs comprising $G_i$, which occur on level i+2 of the tree. Our main task will be to route wires for those edges that have precisely one end-point in a subgraph. Our strategy will be to route wires for such edges in a bottom-up manner, and to connect up these edges only when we process the level of the decomposition tree where these edges were removed. To aid the reader in following this procedure, we include Figs. 3 and 4.

Let us concentrate on one graph $G_i$ at level i of G's decomposition tree, and on the four subgraphs comprising $G_i$ at level i+2 of the tree. Assume inductively that we have at hand one-active-layer layouts for these four subgraphs, each layout having height $H_{i+2}$ and a square base of side

$$S_{i+2} =_{def} h \frac{F \log(N/F)}{H} 2^{-(i+2)/2}$$

(h being the constant in the statement of the theorem). [Since $S_{i+2} \geq (N/2^{i+2})^{1/2}$ when H is in the indicated range, the base of the layout is indeed big enough to accommodate one-fourth of $G_i$'s vertices.] Assume further that each edge of G that has precisely one end in one of the subgraphs, is represented by a wire routed from the appropriate vertex of that subgraph to the top layer of the layout. Finally, assume that each one of these "dangling" edges terminates at a *port* at the top of the layout and that these ports are evenly distributed across the top layer of the layout. (By a *port* here we mean an end of a wire that can be extended upwards if additional layers are added to the top of the embedding; our assumption about even distribution means that the ports are spaced uniformly, as suggested in Fig. 3.) We show now how to construct from the layouts of these level-(i+2) subgraphs of $G_i$ an inductively consistent layout of $G_i$, having height

$$H_i = H_{i+2} + \frac{H}{d \log(N/F)}$$

for some suitably large constant d.

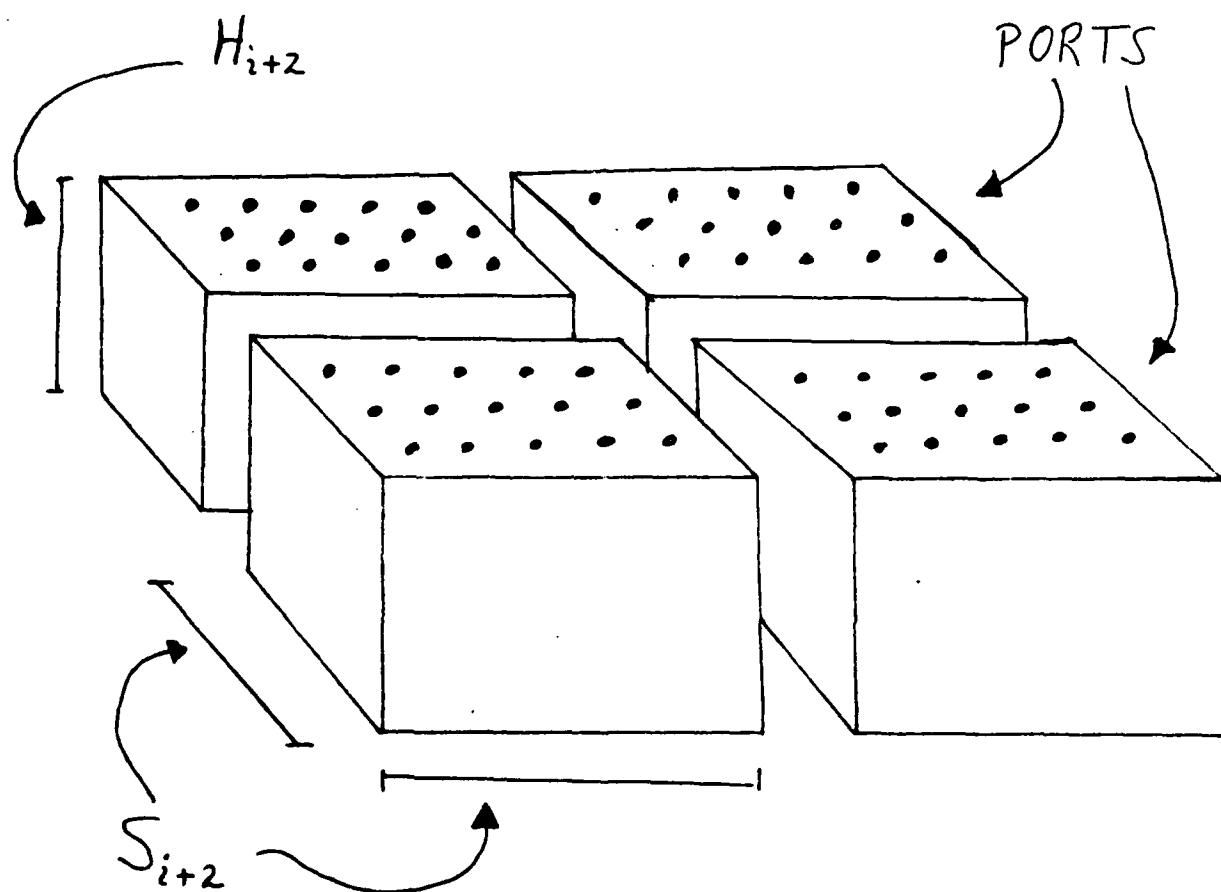We begin by merging the layouts of the four subgraphs into a single "box" having

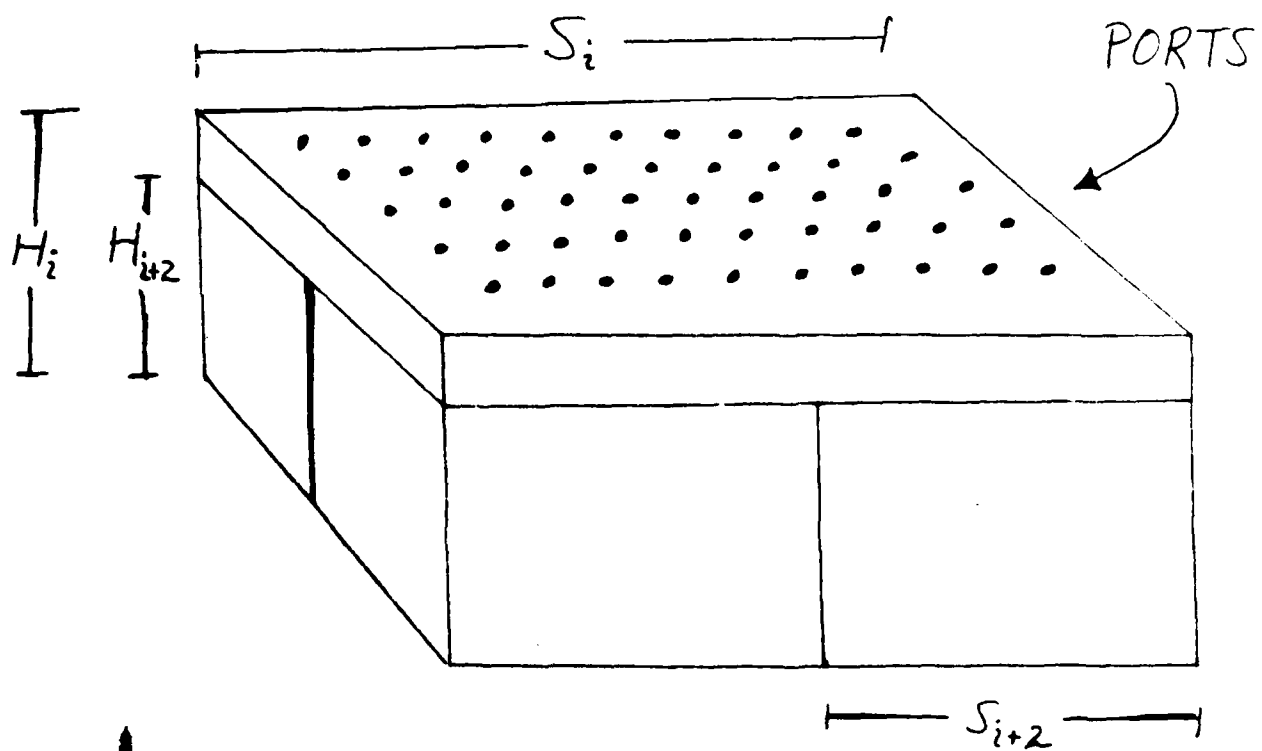**Figure 3.** The given one-active-layer layouts of the four subgraphs of $G_i$.

**Figure 4.** Coalescing the one-active-layer layouts of the subgraphs of $G_i$.

height $H_{i+2}$ and having a square base of side

$$S_i =_{def} h \frac{F \log(N/F)}{H} 2^{-i/2}.$$

We then add $H/(d \log(N/F))$ new (empty) layers to the top of the box, thereby building it up to height $H_i$; see Fig. 4. Next we establish the ports at the top of the new box, that will be needed to extend this construction to a yet-higher level of G's decomposition tree. By Lemma 2.2, no more than $bF/2^{i/2}$ edges of G have precisely one end in $G_i$, for some specified constant b; hence we need create at most this many ports at the top of the new box. We create these ports, spaced evenly throughout the top layer. Finally, we are ready to turn to the task of routing the wires incident to the ports of the original boxes (in layer $H_{i+2}$). Some of these wires must get routed to other ports in the same layer and some to the new ports at the top of the new box.

We effect the necessary routings by using each of the new layers to route an average of $S_i/2$ wires to their appropriate row and column (in one of the new layers). Final connections will then be a simple matter. Since we need to route at most $\frac{3}{2}bF/2^{i/2}$ wires in all, the allotted number of empty layers (namely, $H/(d \log(N/F))$) will suffice, provided that h was chosen sufficiently large. We now describe the details of the routing.

*Layer Assignment*. The first phase of the routing assigns each wire to the layer of the embedding on which it will be routed. To this end, we temporarily superimpose layers $H_{i+2}$ and $H_i$ of the embedding; and we partition the resulting pseudo-layer into $S_i/4$ square regions of area $4S_i$ each. Let M denote the multigraph which has one vertex corresponding to each of these square regions and one edge linking vertices $R_x$ and $R_y$ of M for each wire that must be run in the embedding to connect a port of the square region $R_x$ with a port of the square region $R_y$. Since the number of ports per unit area in the pseudo-layer is at most

$$3b \frac{F/2^{i/2}}{S_i^2} = \frac{3b}{s^2} \frac{2^{i/2}H^2}{F \log^2(N/F)}.$$

the maximum vertex-degree of M does not exceed

$$D = \frac{12b}{h} \frac{H}{\log(N/F)}$$

(= the number of ports per square region). Recall that (by Lemma 2.3) M can be edge-colored using at most $3D/2$ colors. Therefore, provided only that the constant h is chosen sufficiently large (h > log bd suffices), it is now an easy matter to allocate wires to layers: we use each layer $H_{i+2}+k$ of the embedding to route all wires that correspond to edges of M that received the color k.

*Intra-layer Routing*. The second phase of the routing gets each wire to the appropriate row and column of its assigned layer. This is a two-dimensional problem consisting of routing $S_i/4$ wires in a square grid of side $S_i$. In the absence of further information, this might be an impossible task, since the endpoints of the wires to be routed might be configured in a way that did not afford enough room to route the wires. In our case, however, we have distributed the wires' endpoints sufficiently sparsely that the routing is guaranteed to be possible: at most four wires terminate in each square region of area $4S_i$. We have the luxury, therefore, to assign dedicated rows and columns to the wires to be routed. We leave to the reader the details of verifying that this phase of the routing can be accomplished. Note that this routing phase completes the processing of wires that correspond to edges of $G_i$: all connections are made at one of the levels $H_{i+2}+k$.

*Port Connections*. The final phase of the routing connects those wires that correspond to edges having an endpoint outside of $G_i$ to one of the ports at level $H_i$. This, however, is a triviality, since wires are already in the appropriate row and column, and there is no contention for the interlayer route that the wire must traverse.

*The Costs of the Layout*. It remains to assess the efficiency of our embedding. If we apply Lemma 2.2 carefully (as do Bhatt and Leighton [2,13] when treating two-dimensional layouts), we find that we can always force the edges in a fully balanced

decomposition tree of a graph having an $(F, 2^{1/2})$-bifurcator to stay in the top $c$ $\log(N/F)$ levels of the tree, for some appropriate constant $c$. We find thereby that if we have chosen the constant d in the height-recurrence judiciously (it suffices that $d>c$), then the recurrence for the height $H_0$ of the final layout solves to $H_0 \leq H$. The area of the base of the layout never changes throughout the construction: it is always

$$S_0^2 = [h \frac{F}{H} \log(N/F)]^2.$$

Since the Volume of the layout is just $H_0 S_0^2$, we have established the claimed upper bound on $VOL_{1-AL}(G;H)$.

With regard to Wire-Length, it is straightforward to verify that the longest path an uninterrupted wire is stretched over, is proportional to the sum of the linear dimensions of the layout, i.e., $(const)(S_0+H_0)$, whence the claimed bound.

Finally, we remark that no appreciable further decrease in Volume can be obtained by further increasing the height: when H assumes its maximal value, the area of the base of the layout is just some small constant multiple of N. Since all of G's N vertices must reside on a single layer, this area can not decrease further, so subsequent increases in the height can only increase the Volume.

*Computation Time.* The only part of the described layout procedure that is not clearly doable in polynomial time is the generation of an $(F, 2^{1/2})$-decomposition tree for G. And, we assume that we are given such a tree as input to the layout procedure. As an equally efficient alternative to our being given the decomposition tree, we could be given a two-dimensional layout for G as our starting point. We expand on this momentarily. []

Theorem 3.1 affords us the following strengthened version of Rosenberg's [24] results about arbitrary graphs.

*Corollary 3.2.* For any N-vertex graph G and any height $H \leq hN^{1/2}$,

$$VOL_{1-AL}(G;H) \leq (const) \frac{N^2}{H}$$

-19-

and

$$WL_{1-AL}(G;H) \le (const) \frac{N}{H}.$$

At most $(const)N^{1/2}$ layers are needed to minimize $VOL_{1-AL}$. Constructions achieving these results can always be found in time polynomial in $N$.

*Proof.* The worst case in Theorem 3.1 is when $F = N/2$, whence the claimed bounds. In this case, the bifurcator is trivial, and the recursive construction has just one level. []

By judiciously combining two-dimensional layout results with Theorem 3.1, it is not difficult to derive the following AREA-$VOL_{1-AL}$ tradeoff.

*Theorem 3.3: THE ONE-ACTIVE-LAYER AREA-VOLUME TRADEOFF.*

Let $G$ be an $N$-vertex graph, and let $A = AREA(G)$.

*Height-H Layouts.*

There is a constant $h>0$ such that for any height $H$ in the range

$$1 \le H \le h\left(\frac{A}{N}\right)^{1/2} \log(N^2/A),$$

$$\max((NA)^{1/2}, \frac{A}{H}) \le VOL_{1-AL}(G;H) \le (const)\frac{A}{H}\log^2(N^2/A);$$

and

$$(const)\max\left[\frac{A^{1/2}}{N^{1/2}\log(N^2/A)}, \frac{A}{HN\log^2(N^2/A)}\right] \le WL_{1-AL}(G;H) \le (const)\frac{A^{1/2}}{H}\log(N^2/A).$$

*Unrestricted-Height Layouts.*

$$(NA)^{1/2} \le VOL_{1-AL}(G) \le (const)(NA)^{1/2}\log(N^2/A)$$

and

$$(const)\frac{A^{1/2}}{N^{1/2}\log(N^2/A)} \le WL_{1-AL}(G) \le (const)N^{1/2}.$$

Moreover, the value of $H$ that minimizes $VOL_{1-AL}$ is at most

$$(const)\left(\frac{A}{N}\right)^{1/2}\log(N^2/A).$$

Finally, given the Area-A layout of $G$, the embeddings yielding the upper bounds can be found in time polynomial in $N$.

*Proof.*

*The Lower Bounds.*

The lower bounds follow from the lower bound arguments of Theorem 3.1 and the fact [2,13] that $A \le F^2 \log^2(N/A)$.

*The Upper Bounds.*

As in Theorem 3.1, we can establish both the restricted-height and many-active-layer-height upper bounds simultaneously, since the latter bound follows from the former by merely plugging in the maximum permissible value for H. We obtain the restricted-height upper bound in stages. First we recall from Theorem 3.1 that

$$VOL_{1-AL}(G;H) \le (\text{const})\frac{F^2}{H}\log^2(N/F).$$

and

$$WL_{1-AL}(G;H) \le (\text{const})\frac{F}{H}\log(N/F).$$

We next note from Theorem 2.1 that

$$F^2 \le A.$$

Finally, we claim that $1/F < N/A$ so that

$$\log(N/F) < \log(N^2/A),$$

which completes the proof of the upper bound. This final claim is the culmination of the following sequence of inequalities, each following from its predecessors and/or Theorem 2.1.

For $x>1$, $x < 2^x$; hence, $\log(N/F) < N/F$, so $F \log(N/F) < N$. By Theorem 2.1, then, $A < FN$, whence the claim.

The efficiency of actually computing the embeddings that yield the upper bounds follows as in Theorem 3.1, once one performs a recursive decomposition of G by cutting the two-dimensional layout recursively along the lines of the proof of the lower bound in Theorem 3.1. []

-21-

It is worth noting that the upper bounds in Theorem 3.1 are everywhere existentially tight (within constant multiples) for every value of N, F, and H; i.e., the factors of $\log(N/F)$ cannot be avoided. To verify this, one needs recall that Leighton [13] proved that the upper bounds in Theorem 2.1 are everywhere existentially tight:

For all N and F, there exist N-vertex graphs G whose smallest $2^{1/2}$-bifurcators have size F such that

$$c_1[F \log(N/F)]^2 \le AREA(G) < c_2[F \log(N/F)]^2.$$

For any one of these maximal-AREA graphs G, the lower bounds of Theorem 3.3 assure us that $VOL_{1-AL}(G;H)$ is no smaller than some constant multiple of

$$\max(N^{1/2}F \log(N/F), \frac{F^2}{H}\log^2(N/F)).$$

This information combines with the upper bounds of Theorem 3.1 to establish the claimed tightness. We do not know that the upper bounds of Theorem 3.3 are similarly tight, and, indeed, we conjecture that they are not.

*Conjecture 3.4: THE ONE-ACTIVE-LAYER AREA-VOLUME TRADEOFF*

Let G be an N-vertex graph, and let $A = AREA(G)$. There is a constant h>0 such that, for any height H in the range

$$1 \le H \le h(AN)^{1/2},$$

we have

$$VOL_{1-AL}(G;H) = (const)\frac{A}{H}.$$

For larger H, no additional decrease in volume can be obtained.

**3.2. Unrestricted Layouts.**

We turn now to the task of proving analogs of Theorems 3.1 and 3.3 for many-active-layer three-dimensional layouts. We shall be less thorough in our pursuit of the analogs, for the following reasons.

-22-

1. The major ideas required to obtain compact three-dimensional embeddings appear already in the one-active-layer case, which we have looked at in great detail.

2. While the one-active-layer case can already be considered to have been realized (say, in IBM's TCM [10,22]), it is not yet clear that many-active-layer three-dimensional layouts will ever be more than a theoretical construct.

3. The many-active-layer results that we develop suggest that only relatively minor gains are achieved by abjuring the one-active-layer restriction.

4. The technical details of obtaining height-restricted many-active-layer layouts are substantial and may not be worth the gains over the one-active-layer case.

The major change in layout strategy in the many-active-layer model is that we must use $2^{2/3}$-bifurcators in order to obtain compact layouts. Indeed, the feature that renders restricted-height layouts prohibitively complicated with the many-active-layer model is that one must play off $2^{1/2}$-bifurcators against $2^{2/3}$-bifurcators. We satisfy ourselves, therefore, with the following many-active-layer results.

*Theorem 3.5: THE UNRESTRICTED THREE-DIMENSIONAL LAYOUT THEOREM*

Let G be an N-vertex graph, let F be the size of its minimum $2^{2/3}$-bifurcator, and let A = AREA(G). The many-active-layer three-dimensional layouts of G satisfy

$$F^{3/2} \leq VOL(G) \leq (const)[F \log(N/F)]^{3/2}.$$

and

$$(const)\frac{F^{5/2}}{N^2 \log^{1/2}(N/F)} \leq WL_3(G) \leq (const)[F \log(N/F)]^{1/2}.$$

*Proof.* We establish the lower and upper bounds in turn.

*The Lower Bounds.*

Let G be laid out in an H×W×L grid, where with no loss of generality, H≤W≤L.

We establish the lower bound on Volume by recursively bisecting the layout of G, much as we did in the lower-bound proof of Theorem 3.1. We slice the H×W×L grid

-23-

holding the layout into two H×W×(L/2) grids, purposely choosing to bisect the biggest of the dimensions. We then recursively continue this bisecting, each time halving the biggest dimension of the grid being sliced. Now, at stage i of this bisecting, we are slicing boxes of volume $HWL/2^i$ (we started at stage 0). When slicing each such box, the plane of the slice has area at most $(HWL/2^i)^{2/3}$ -- the longest dimension is at least $(HWL/2^i)^{1/3}$, leaving only the indicated area for the plane. Since wires have all unitcross-sections, each slice cuts no more than $(HWL/2^i)^{2/3}$ wires. Since G can be so bisected recursively, and since we have been looking at an arbitrary three-dimensional layout of G, it follows that G has a $2^{2/3}$-bifurcator of size $(VOL(G))^{2/3}$. Since this bifurcator is at least as big as G's minimum such bifurcator F, we have thus shown that

$$F^{3/2} \le VOL(G).$$

The lower bound on wire-length depends on the fact that any graph with a $2^{1/2}$-bifurcator of size $\overline{F}$ has a $2^{2/3}$-bifurcator of size

$$F \le N^{1/3}\overline{F}^{2/3}.$$

To verify this, we need only check that

$$\min(\frac{N}{2^i}, \frac{\overline{F}}{2^{i/2}}) \le \frac{N^{1/3}\overline{F}^{2/3}}{2^{2i/3}}$$

for all i. (Recall that at most $\min(N/2^i, \overline{F}/2^{i/2})$ edges are cut in any i-th level partition of a $2^{1/2}$-bifurcator.) This inequality is easily verified since

$$\frac{N}{2^i} \le \frac{N^{1/3}\overline{F}^{2/3}}{2^{2i/3}}$$

for $i \ge 2\log(N/\overline{F})$, and $\dfrac{\overline{F}}{2^{i/2}} \le \dfrac{N^{1/3}\overline{F}^{2/3}}{2^{2i/3}}$

for $i \le 2\log(N/\overline{F})$.

By inverting the preceding inequality, one can show that the smallest $2^{1/2}$-bifurcator $\overline{F}$ of any N-vertex graph G is at least

$$\frac{F^{3/2}}{N^{1/2}}$$

where F is the size of the graph's smallest $2^{2/3}$-bifurcator. We can now invoke the arguments of Theorem 3.1 to conclude that

$$WL_3(G) \geq \frac{WL_2(G)}{(VOL(G))^{1/3}}$$

$$\geq (const) \frac{\overline{F}^3}{N(VOL(G))^{1/3}}$$

$$\geq (const) \frac{F^3}{N^2[F \log(N/F)]^{1/2}}$$

$$= (const) \frac{F^{5/2}}{N^2 \log^{1/2}(N/F)}.$$

which is the claimed lower bound.

*The Upper Bounds.*

As in Theorem 3.1, the upper bounds here are significantly more complicated to establish than the lower bounds. Once again, our proof is via an explicit inductive construction. In this case, the construction will lay out a subgraph $G_i$ of G that resides at level i of a $(2^{2/3}$-)decomposition tree of G, by combining the *layouts of the eight* subgraphs of $G_i$ that reside at level i+3 of the tree.

Let us assume: (1) that we are given layouts of the eight level-(i+3) subgraphs, in "boxes" of dimensions $H_{i+3} \times W_{i+3} \times L_{i+3}$ each, (2) that the boxes are all similarly oriented in space (so we can talk about their fronts and tops, etc.), and (3) that each of the boxes has on its front face a *connection grill* which is a set of

$$b \frac{F_{i+3}}{H_{i+3}}$$

length-$H_{i+3}$ columns, spread evenly across the width-$W_{i+3}$ front face, collectively comprising the

$$F_{i+3} =_{def} \frac{F}{(2^{2/3})^{i+3}}$$

ports through which we shall wire these subgraphs to each other and to the remainder of G. See Fig. 5.

As the first step in laying out $G_i$, we place our eight boxes in a big box of dimen-
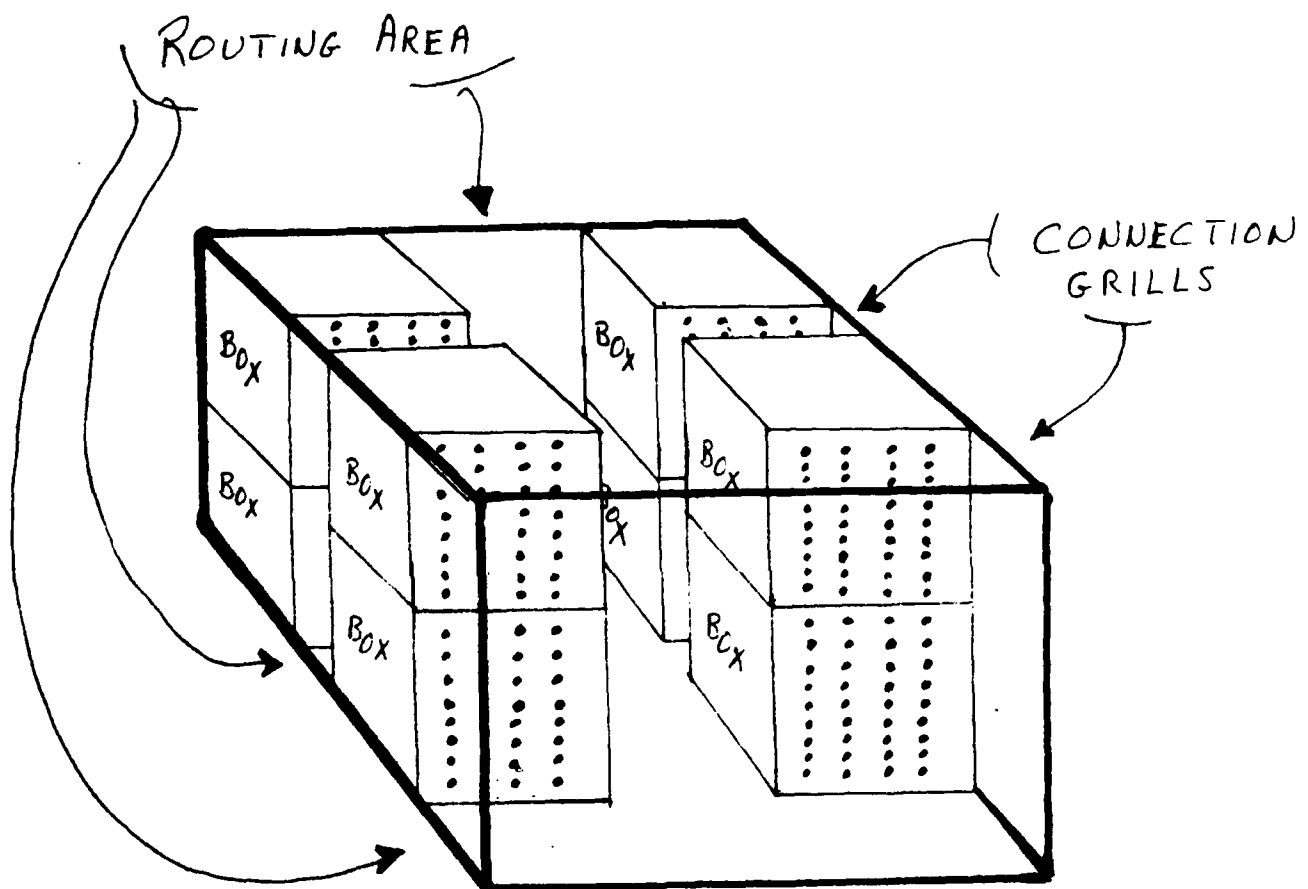
-25-

**Figure 5.** Coalescing the many-active-layer layouts of the subgraphs of $G_i$.

sions $H_i \times W_i \times L_i$, where

$$H_i = 2H_{i+3}$$

$$W_i = 2W_{i+3} + 2b\frac{F_{i+3}}{H_{i+3}}$$

$$L_i = 2L_{i+3} + 4b\frac{F_{i+3}}{H_{i+3}} + c\frac{F_{i+3}}{W_i}.$$

for appropriately chosen constants b,c. We place the small boxes as follows. Four of the boxes are placed at the four corners of the back of the big box. In front of these boxes we place $bF_{i+3}/H_{i+3}$ empty layers that will be used for wire-routing. In front of these empty layers, we place the four remaining small boxes, in the corners. In front of these boxes we place $3bF_{i+3}/H_{i+3} + cF_{i+3}/W_i$ empty layers for wire-routing. We complete the placement by placing $2bF_{i+3}/H_{i+3}$ empty layers between the left and right tiers of small boxes. Finally, on the front face of the big box, we uniformly spread out a new connection grill with $bF_i/H_i$ columns of $H_i$ ports each, containing the $bF_i$ ports that are guaranteed to be sufficient to wire $G_i$ up to the rest of G. See Fig. 5.

Now we turn to the task of routing the wires that leave the level-(i+3) graphs (through their connection grills), both among each other and to the new connection grill. As the first step of this routing, we use the $bF_{i+3}/H_{i+3}$ empty layers in front of the four rear boxes to route the wires from these boxes' connection grills into one big $[H_i \times 2(bF_{i+3}/H_{i+3})]$ rectangular *connection patch* in between and in front of the four rear boxes (a grill is spread out, while a patch is compressed); this can be accomplished by having the innermost columns of the grills move in on one layer to meet one another, the next-to-innermost move in on the next layer to be adjacent to the innermost ones, and so on; see Fig. 6. The next step is to run the connection patch, which occupies the $2bF_{i+3}/H_{i+3}$ routing layers between the left and right tiers of boxes, to the empty layers that we have placed front of all the boxes (also depicted in Fig. 6). Now we take the big connection patch we have just run from the back of the layout,
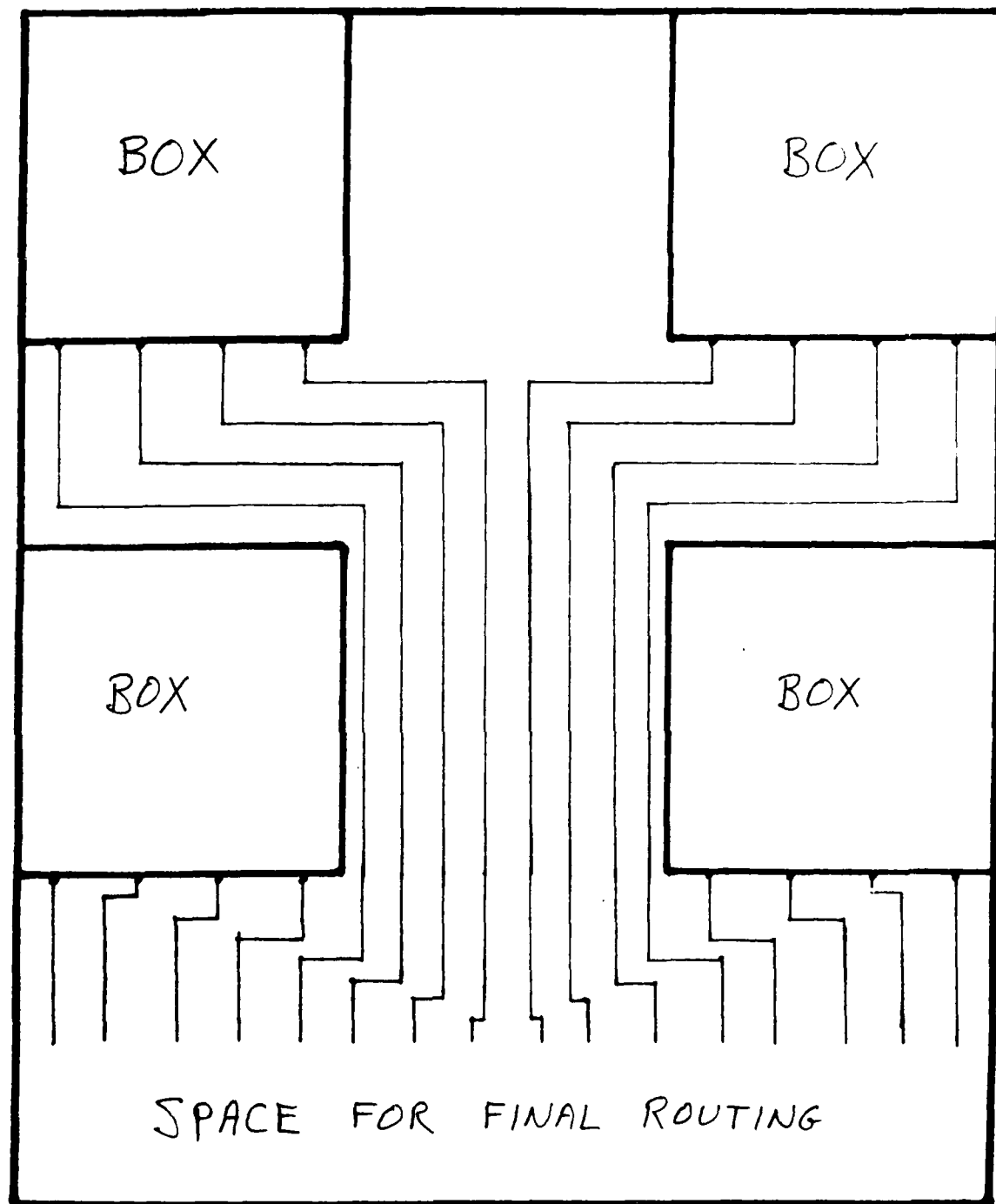
**Figure 6.** A view from above of the coalescing of the wires from the subgraphs.

and the small connection grills on the front tier of small boxes, and we use $2bF_{i+3}/H_{i+3}$ of the empty layers at the front of the box to distribute these columns of ports so that they are evenly spaced across the front of the layout, i.e., so that they become a connection grill. (There are $4bF_{i+3}/H_{i+3}$ columns to be distributed across a face of width $W_i$, so the columns are spaced with

$$\frac{W_i H_{i+3}}{4bF_{i+3}}$$

empty columns intervening between full columns.) As the next-to-final step, we assign the ports of this new connection grill to the layers on which they will be routed to their appropriate rows and columns. This assignment is done using the same device as in Theorem 3.1. We partition the face with the ports into rectangles of height $c_1 F_i/W_i$ and width $(W_i H_{i+3})/(4bF_{i+3})$. By design, each of these rectangles contains no more than $c_1 F_i/W_i$ ports. Hence, when we follow our ploy from Theorem 3.1 and form the multigraph M from the partition, and edge-color M, we are assured by Lemma 2.3 that we need use no more than $(3/2)c_1 F_i/W_i$ colors. Hence, when we assign wires to layers by their colors, as in Theorem 3.1, we need use no more than $(3/2)c_1 F_i/W_i$ layers. As in Theorem 3.1, it is an easy matter from this point, to complete the routing using the as-yet-unused reserved routing layers, once having assigned layers. We leave details to the reader.

It remains to assess the efficiency of the layouts produced by the preceding construction. If we arbitrarily set

$$H = H_0 = h[F \log(N/F)]^{1/2},$$

which is acceptable providing that the constant h is chosen judiciously, then we find that the recurrences for length and width solve to

$$W = W_0 = (\text{const})[F \log(N/F)]^{1/2},$$

and

$$L = L_0 = (\text{const})[F \log(N/F)]^{1/2}.$$

(In solving these recurrences, one must keep in mind that F is the size of G's smallest

$2^{2/3}$-bifurcator.) The claimed upper bound on Volume follows by just multiplying these linear dimensions; the upper bound on Wire-Length follows by summing them, for our layout scheme requires a wire to traverse each linear dimension only a small number of times. []

The bounds of Theorem 3.5 can be shown to be existentially tight via inheritance from the bounds of [13].

As with one-active-layer embeddings, we can derive from our general layout theorem an Area-Volume tradeoff.

*Theorem 3.6: THE GENERAL AREA-VOLUME TRADEOFF.*

Let G be an N-vertex graph, and let A = AREA(G). The many-active-layer three-dimensional layouts of G satisfy

$$\max(N, A^{3/4}) \leq \text{VOL}(G) \leq (\text{const})(AN)^{1/2}.$$

and

$$\text{WL}_3(G) \leq (\text{const})(AN)^{1/6}.$$

*Proof.*

*The Lower Bounds.*

The lower bounds follow by previously enunciated principles: the vertices of G alone, ignoring wires, consume volume N; and any H×L×W layout of G (where H ≤ L,W) can be transformed into an LH×WH two-dimensional layout for G, whence AREA(G) ≤ VOL(G)$^{4/3}$.

*The Upper Bounds.*

The upper bounds follow from the construction of Theorem 3.5 and the fact that any graph admitting an Area-A layout has a $2^{1/2}$-bifurcator of size $A^{1/2}$, hence a $2^{2/3}$-bifurcator of size $(NA)^{1/3}$ -- cf. the proof of Theorem 3.5. If we plug this $2^{2/3}$-bifurcator into the upper bound of Theorem 3.5, we find that

$$\text{VOL}(G) \leq (\text{const})(NA)^{1/2}\log^{3/2}(N^2/A)$$

and

$$WL_3(G) \leq (\text{const})(NA)^{1/6}\log^{1/2}(N^2/A)$$

A careful analysis of the layout of G hidden in this upper bound indicates that we are really cutting more edges of G at each step than a smaller $2^{2/3}$-bifurcator of G would force us to (since our bound on F is very conservative). Hence, when we calculate in detail the dimensions of the layout produced with this big bifurcator, we find that we actually avoid the logarithmic factor, and we obtain the bounds of the Theorem. []

## 4. CONCLUSIONS.

The work in this paper leaves the reader with a number of messages, which we now encapsulate.

* Three-dimensional layouts can be appreciably more conservative of resources, both material and wire-length, than can two-dimensional layouts.

* Three-dimensional layouts are not appreciably harder to "compute" than are two-dimensional layouts; in fact the former can be produced from the latter algorithmically.

* For many classes of graphs, one-active-layer three-dimensional layouts are as efficient as many-active-layer layouts. In general the best bounds that can be proved in terms of A (optimal two-dimensional Area) and N (number of vertices) for the two classes of layouts differ by at most a logarithmic factor. Thus no general layout procedure is uniformly better than our one-active-layer layout procedure. If this phenomenon occurs also in practice, then the value of fabricating transistors on multiple levels is limited.

* Even in the one-active-layer model, only a limited numbers of layers are helpful. Roughly speaking, only $(A/N)^{1/2}$ or $F/N^{1/2}$ layers lead to increased efficency of multilayer layouts: additional layers cannot further decrease Volume. It is worth
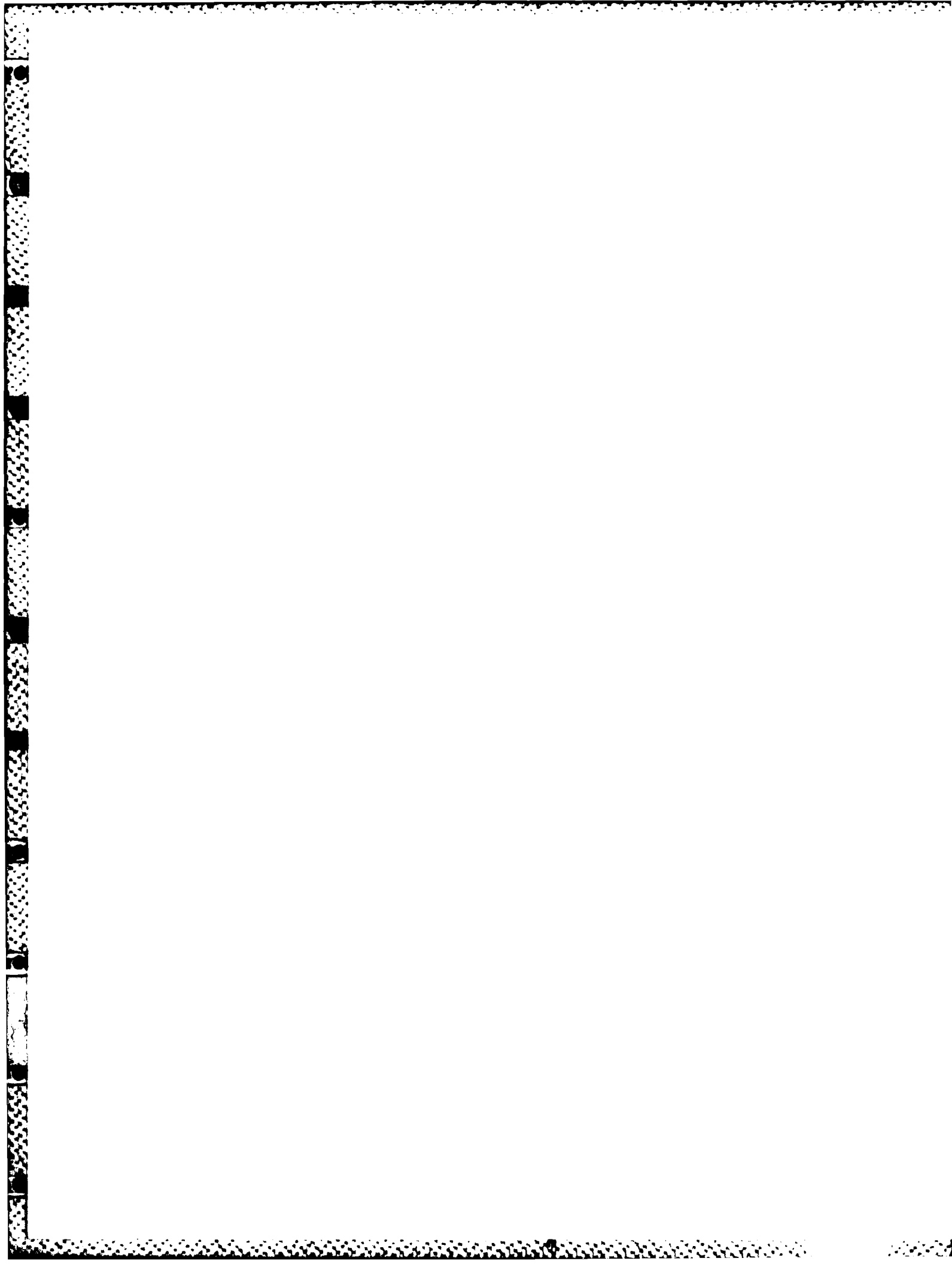
noting that the quantity $(A/N)^{1/2}$ is closely related to the complexity of two-dimensional placement and routing: this is the average channel width in a two-dimensional layout of the circuit. Although we did not prove it in this paper, we suspect that additional layers are also not useful in decreasing wire-length. We know this to be the case for many families of graphs.

* Although we have not paid attention to issues like the size of constants here, it seems likely that the method of layer assignment that we employed in the upper-bound proof of Theorem 3.1 can be adapted to produced computationally efficient assignments in practical situations.

## REFERENCES

1. V.E. Benes (1964): Optimal rearrangeable multistage connecting networks. *Bell Syst. Tech. J. 43*, 1641-1656.

2. S.N. Bhatt and F.T. Leighton (1984): A framework for solving VLSI graph layout problems. *J. Comp. Syst. Sci.*, to appear.

3. M.L. Brady and D.J. Brown (1984): Arbitrary planar routing with four layers. *1984 MIT Conf. on Advanced Research in VLSI*, 194-201.

4. R.D. Etchells, A.D. Cummings, J. Grinberg, G.R. Nudd (1982a): Cellular array architecture for microelectronic implementation. In preparation.

5. R.D. Etchells, A.D. Cummings, J. Grinberg, G.R. Nudd (1982b): A yield-redundancy policy for wafer-scale integration. In preparation.

6. R.D. Etchells, A.D. Cummings, J. Grinberg, G.R. Nudd (1982c): Power dissipation in 3-D VLSI. In preparation.

7. R.D. Etchells, J. Grinberg, G.R. Nudd (1981): Development of a three-dimensional circuit integration technology and computer architecture. *Soc. Photographic and Instrumentation Engineers 282*, 64-72.

8. J.F. Gibbons (1982): SOI -- a candidate for VLSI? *VLSI Design III*, 54-55.

9. L.S. Heath (1983): Multilayer channel routing. MCNC Tech. Rpt. 83-06.

10. C.W. Ho (1982): High performance VLSI computer packaging. *1982 MIT Conf. on Advanced Research in VLSI*, p. 42.

11. Hughes Research Laboratories (1982): A cellular VLSI architecture for image analysis and two-dimensional signal processing. Typescript.

12. H.W. Lam, A.F. Tasch, Jr., T.C. Holloway (1980): Characteristics of MOSFETs fabricated in laser-recrystallized polysilicon islands with a retaining wall structure on an insulating substrate. *Electron Dev. Let. EDL-1*, (1980), 206-208.

13. F.T. Leighton (1982): A layout strategy for VLSI which is provably good. *14th ACM Symp. on Theory of Computing*, 85-98.

14. F.T. Leighton (1983): *Complexity Issues in VLSI: Optimal Layouts for the Shuffle-Exchange Graphs and other Networks*. MIT Press, Cambridge, MA.

15. F.T. Leighton and A.L. Rosenberg (1983): Automatic generation of three-dimensional circuit layouts. *1983 IEEE Intl. Conf. on Computer Design: VLSI in Computers*, 633-636.

16. C.E. Leiserson (1983): *Area-Efficient VLSI Computation*. MIT Press, Cambridge, MA.

17. D.B. Lenat, W.R. Sutherland, J. Gibbons (1982): Heuristic search for new microcircuit structures: An application of artificial intelligence. *The AI Magazine 3*, 17-33.

18. W.R. Locke (1983): Three-dimensional integration: A critical survey. MCNC Tech. Rpt. 83-08.

19. E.W. Maby and D.A. Antoniadis (1981): Device structures for three-dimensional integration. Lecture at MIT Research Review, December, 1981.

20. F.P. Preparata (1981): Optimal three-dimensional VLSI layouts. *Math. Systems Theory 16*, 1-8.

21. F.P. Preparata and W. Lipski (1982): Optimal three-layer channel routing. Tech. Rpt. ACT-34, Coordinated Science Lab., Univ. of Illinois.

22. D. Rosen (1981): TCM - it's a new word for density in logic packaging. *THINK*, p. 23.

23. A.L. Rosenberg (1981): Three-dimensional integrated circuitry. In *VLSI Systems and Computations* (ed. H.T. Kung, B. Sproull, G. Steele) Computer Science Press, Rockville, MD, pp. 69-80.

24. A.L. Rosenberg (1983): Three-dimensional VLSI: A case study. *J. ACM 30*, 397-416.

25. C.E. Shannon (1949): A theorem on coloring the lines of a network. *J. Math. Physics 28*, 148-151.

26. C.D. Thompson (1980): A complexity theory for VLSI. Ph.D. Thesis, Carnegie-Mellon University; see also Area-time complexity for VLSI. *11th ACM Symp. on Theory of Computing*, 1979, 81-88.

27. L.G. Valiant (1981): Universality considerations in VLSI circuits. *IEEE Trans. Comp., C-30*, 135-140.

28. Z.A. Weinberg (1981): Polysilicon recrystallization by $CO_2$ laser heating of $SiO_2$. IBM Report RC-8835.

29. D.S. Wise (1981): Compact layouts of banyan/FFT networks. *VLSI Systems and Computations* (ed. H.T. Kung, B. Sproull, G. Steele) Computer Science Press, Rockville, MD, pp. 186-195.

OFFICIAL DISTRIBUTION LIST

1984

Director                                                2 Copies
Information Processing Techniques Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA   22209


Office of Naval Research                                 2 Copies
800 North Quincy Street
Arlington, VA   22217
Attn:  Dr. R. Grafton, Code 433


Director, Code 2627                                      6 Copies
Naval Research Laboratory
Washington, DC   20375


Defense Technical Information Center                     12 Copies
Cameron Station
Alexandria, VA   22314


National Science Foundation                             2 Copies
Office of Computing Activities
1800 G. Street, N.W.
Washington, DC   20550
Attn:  Program Director


Dr. E.B. Royce, Code 38                                 1 Copy
Head, Research Department
Naval Weapons Center
China Lake, CA   93555


Dr. G. Hopper, USNR                                     1 Copy
NAVDAC-OOH
Department of the Navy
Washington, DC   20374